

## A TWO-LEVEL METHOD TO CONTROL GRANULAR SYNTHESIS

Andrea Valle, Vincenzo Lombardo

MultiLab e Dipartimento di Informatica - Università di Torino  
andrea.valle@unito.it, vincenzo@di.unito.it

### ABSTRACT

This paper presents a formal system for the algorithmic control of composition based on granular synthesis. The system features two description levels: a low level, that organizes grains into a graph structure, and a high level, that distributes the graphs of the low level in specific locations of a space. A composition is a trajectory in the space, appropriately interpreted to control a number of parameters of physical and musical relevance. The paper is organized as follows: first, we introduce the composition process with granular synthesis and we briefly outline the current approaches to control; second, we describe the formal system in terms of the two levels that compose it; finally, we see how the system can be viewed as a generalization of the note approach and the stochastic approach.

### 1. THE COMPOSITION WITH GRANULAR SYNTHESIS

Granular synthesis is a general term that encompasses various kinds of synthesis techniques based on a grain representation of sound, i.e. sonic events are built from “elementary sonic elements” of very short duration ([1]; as a general reference, see [2]). Different organization techniques can lead to very different timbral and compositional results (see [1], [2], [3], [4]). So, one of the main questions arising while working with grains is how to move from single grain level (microstructure) up to compositional design (macrostructure), possibly passing through note level (ministructure) and rhythm level (mesostructure) (following [5]: 266; see also [2]: 3). We can distinguish two major approaches: the *note* approach and the *stochastic* approach.

In the case of the note approach, the focus is on microstructure as embedded in ministructure: so, ministructure defines the sound objects and granularity defines the timbre of each object (i.e. drum roll, rolled phonemes, flutter-tongue, [6]: 56). Granular synthesis and granulation of existing sound objects are methods to create/transform elements at the “note” level. As in traditional composition, there is a logic gap between sound and structure ([7]). This is the approach implemented in grain-based modules of DSP applications ([8], [1], [4]), and in Csound built-in opcodes ([9]).

More radically, in the stochastic approach, granularity is intended as a compositional feature. Having to work

with a pulviscolar matter, composers involved in granular synthesis have often decided to avoid an “instrumental-music approach” to promote textural shaping as a general compositional feature, in order to “unite sound and structure” ([7]: 120). Various stochastic methods and strategies have been used to control grain densities, distribution in frequency spectrum, waveshape in the time course (see the “classic” works by Xenakis, Roads, Truax). In Xenakis ([5]), the sound is thought as an evolving gas structure. The audible field is modelled according to the Fletcher-Munson diagram, which is subdivided in a finite number of cells. Each instant is described through the stochastic activation of certain cells in the diagram (a “screen”) and each screen has a fixed temporal duration. The sound/composition is an *aggregatum* of screens collected in a “book” in “lexicographic” order (as in the series of sections of a tomography). In Truax ([10], [7]), massive sound texture is obtained via the juxtaposition of multiple grain streams (“voices”, like in polyphony): the parameters of each grain stream are controlled through tendency masks representing variations over time (i.e. timbre selection, frequency range, temporal density, [7], [10]). This approach is well known in the literature as Quasi-Synchronous Granular Synthesis (QSGS, [11], [4], [2]). In Roads [11], grains are scattered probabilistically over frequency/time plain regions (“clouds”). The compositional work relies on controlling cloud global parameters (i.e. start time and duration of the cloud, grain duration, density of grains, etc.). In these three cases, compositional strategies are based on the direct control of the creative process with an empty uniform time/frequency canvas<sup>1</sup>. Not surprisingly, the compositional metaphor in Roads is explicitly related to painting, using different brushes with different (sound) colours ([11]: 143).

The goal of this paper is to provide a new perspective on the composition process with granular synthesis by introducing a formal system based on two description levels. As we see below, the system can be viewed as a generalization of the note and the stochastic approaches.

---

<sup>1</sup> See also the graphic notation in [12]: 156-57. This is the standard spatial metaphor in different granular synthesis implementations using tendency masks: in a Csound-oriented perspective, see for example the software *GSC4* ([13]) and *Cmask* ([14]).

## 2. GEOGRAPHY: A TWO-LEVEL SYSTEM FOR GRAIN GENERATION AND CONTROL STRUCTURE

In this section we describe the formal system GeoGraphy, that models the composition process with two components, one for the generation of grain sequences, another for the parametric control of waveforms.

First, we introduce some terminology. A composition is a set of *tracks*; each track is a grain sequence (Figure 1), where the single grains are waveforms that result from granular synthesis and parametric control. The formal system GeoGraphy consists of two components: a graph-based generator of grain sequences (i.e. tracks), and a map-based controller of grain waveform parameters.

The grain generator (level I) is based on directed graphs, actually a multigraph ([15]), as it is possible to have more than one edge between two vertices (Figure 2). A vertex represents a grain; an edge represents the sequencing relation between two grains. Grains can be either sampled waveforms with fixed durations, or waveforms generated by a synthesis process with a duration that is overtly marked on the vertex (all the vertices in Figure 2 represent sampled grains). A label on an edge represents the temporal distance between the onset times of the two grains connected by the edge itself<sup>2</sup>. A grain sequence is a path on the graph, that in case a graph contains loops can also be infinite. The generation of a grain sequence is achieved through the insertion of dynamic elements into the graph, called *graph actants*. A graph actant is initially associated with a vertex (that becomes the origin of a path); then the actant navigates the graph by following the directed edges according to some probability distribution<sup>3</sup>. Multiple independent graph actants can navigate a graph structure at the same time, thus producing more than one grain sequence.

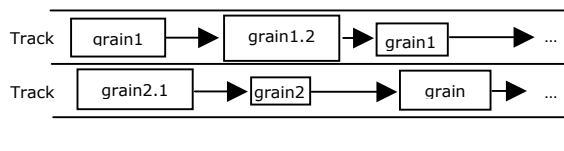


Figure 1: A musical piece

<sup>2</sup> All durations in the formalism can be made dependent on some probability distribution, so to act as a general stochastic grain generator. This feature together with the track (or voice) structure of the musical piece allows GeoGraphy to simulate the expressive power of QSGS.

<sup>3</sup> For those readers that are familiar with Petri nets, a graph actant can be viewed as a *token*. The probabilistic control of the token also reminds to stochastic Petri nets.

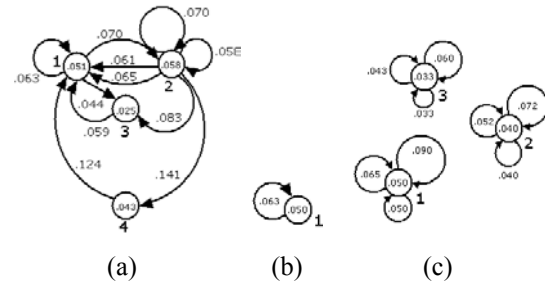


Figure 2: (a) A complex multigraph; (b) A graph of one vertex and one edge; (c) A graph consisting of three disconnected subgraphs.

In Figure 2 there are three examples of graphs. The graph in Figure 2a is a multigraph with several connections (almost completely connected). It also contains loops. One possible result is in Figure 3, where some amplitude control, a typical Gaussian envelope, has been applied to avoid clicking. Starting from vertex 4, the graph actant generates a grain of duration 43 milliseconds (vertex label), then it reaches vertex 1 with a delay of 124 milliseconds (edge label), it loops two times on vertex 1, generating two grains of 51 milliseconds with a delay of 63 milliseconds, then it leaves vertex 1 for vertex 2 and so on. As grain duration and delay of attack time are independent, it is possible to superpose grains (vertex label > edge label, see the last two grains in Figure 3).

In Figure 2b there is a graph with one vertex and one edge that loops on the unique vertex. The grain sequence produced by such a graph is the exact repetition of the grain associated with the vertex; each repetition starts after 63 milliseconds with respect to the beginning of the previous repetition.

In Figure 2c there is a graph consisting of three disconnected subgraphs, each with one vertex and three edges that loop on the vertex itself. If we assume a single actant on each graph, the system generates three simultaneous streams of grains. If we associate each vertex a grain of fixed frequency, we yield a spectrum consisting in three rows (Figure 4), a “stratus” in Roads’ typology ([11]: 165, [2]: 104).

In order to control the setting of the parameters associated with the grain waveforms, the idea implemented in the GeoGraphy system (level II) is to position the graphs in a space, and then to control the parameter values by navigating the space (*control space* or *map of graphs* – Figure 5). Once the single sound streams have been defined through the generation of graphs, the composer distributes the graphs onto a map, and then designs a *trajectory* that allows to decide how the several sound streams contribute to the piece. The control of the parameters occurs with reference to the spatial metaphor: parameters value ranges are mapped onto spatial distance, and the nearer is a trajectory to some vertex, the higher is the value of some parameter for the grain waveform represented by that vertex.

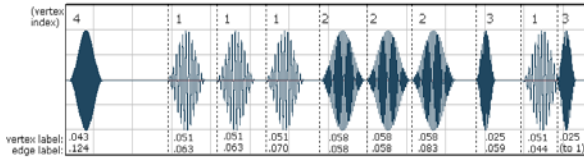


Figure 3: A grain sequence generated by the graph of Figure 2a.

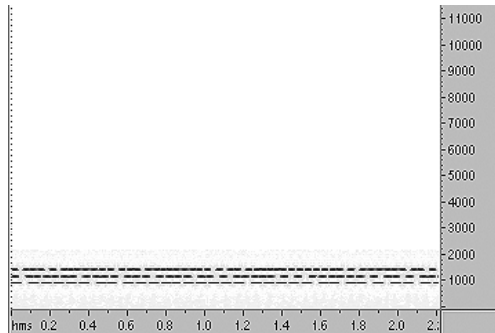


Figure 4: Spectrum of the signal generated by the three subgraphs of Figure 2c ("stratus").

So, the parametric controller of the GeoGraphy system is a spatial *map* of graphs. Theoretically, the space can have any number of dimensions, but we limit our composition model to the Euclidean space (actually, the examples in this paper feature only two dimensions). In order to place the graphs on the map, each vertex is associated with coordinates in the space. A map contains a finite number of graphs ( $n$ ), which work independently, thus generating  $a$  grain sequences, where  $a$  is the total number of the graph actants that navigate in all the graphs. As there is at least one graph actant for each graph, there will be a minimum of  $n$  sequences ( $a \geq n$ ).

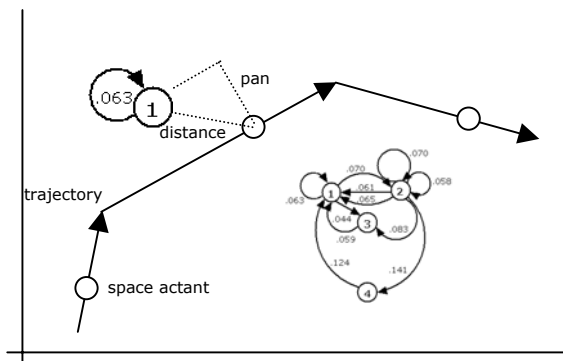


Figure 5: A control space with one space actant and its trajectory, and two graphs. Distance and panning are marked with respect to one vertex

Trajectory design occurs by defining a new actant (the *space actant* – Figure 5) which navigates at constant rate in the space following a *trajectory*. Each vertex emits a grain (determined by the passage of the graph actant); when a grain is generated, the device calculates the distance between the space actant and the vertex. This distance is then used as a general control parameter. If we consider the space actant as a directed human head (Figure 5), the displacement of the vertex from the frontal and the lateral axes is used to control panning, while the Euclidean distance is used to control amplitude. Trajectories can be explicitly defined by the composer or generated algorithmically (e.g., *brown motion*). Different trajectories determines different strategies of exploration of the map space.

The two levels of GeoGraphy are described in detail in ([16]). In the next section we discuss some expressivity issues of the model; in particular we see how it generalizes over the note approach and the stochastic approach.

### 3. EXPRESSIVITY ISSUES

The space typically considered in granular synthesis techniques is the physical continuum time/frequency. This space is not necessarily musically meaningful: it can be quite complex to define musical meaningful parameters while working with it without great efforts. On the contrary, a general map space as the one we have proposed can embed different topologies of musical features among the ones proposed in literature. This can be done through the application of grouping strategies to vertices in the control space. The two compositional dimensions are density (number of grains for surface unit) and structure organization of vertices (i.e. qualitative relationship over grains). For example, in order to enrich the spectral content of a musical object we can collapse several vertices in a single location. The graph structure can contribute to stress the hierarchical relationship in the group: the star-structured graph of Figure 6a generates a sequence of the form  $1n1n...$ , thus stressing the importance of vertex 1. The vertices (or group of vertices) can be distributed following timbral spaces discussed in literature ([17]: 76; [18]; [19]:13; [20]: 59; [21]: 48; [22]: 197), but also pitch spaces like the Two-dimensional Melodic and Harmonic Maps proposed by ([23]: 374-78). Harmonic/inharmonic, sparse/dense, low/high, are categories that, together with topology features like centre/periphery, are of typical use.

If one chooses to consider one of the axis of the map space as frequency, the space actant displacement on that axis can be thought as a band pass filtering (if the distance has a threshold, bandwidth is represented exactly by the diameter of the audible circle, Figure 6b). Vertex groupings in Figure 6c follows cloud patterns as described by ([11]: 166; [4]: 182; [2]: 105) (from left to right: cumulus, stratus, glissandi).

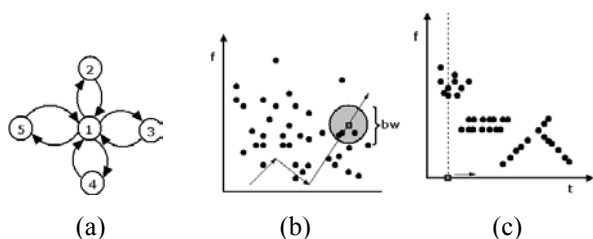


Figure 6: (a) A star graph; (b) Space actant (square) as a bandpass filter (bw: bandwidth) (c) Space actant as a scanning device (audible circle compressed to dashed line) (edges are omitted)

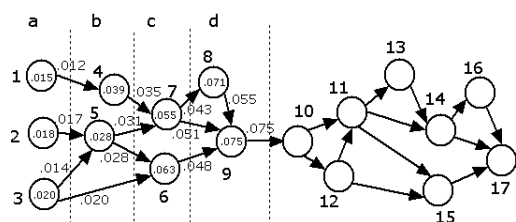


Figure 7: Starting from one of the vertices 1-3 every path has a duration between 143 (vertices 3+6+9) and 221 milliseconds (2+5+7+8+9). The subgraph 1-9 can consequently be considered as a note event leading to the note event, resulting from the subgraph 10-17

Note that the time/frequency domain, and hence the stochastic approaches, can be considered a special case of the map space in which: a) one axis represents frequency; b) the other represents time; c) the trajectory coincides with the time axis; d) the circle surrounding the space actant is compressed to its diameter parallel to the frequency axis; e) there is not distance threshold or, if present, it is as large as the frequency axis.

It must be noted that a map space should be used with caution in simulating a time/frequency space. In fact, the map space is filled not with actual events but only with potential events (vertices), in the sense that it is possible that a grain is emitted at the very moment in which the space actant scanning the timeline trajectory is passing (as the result of the activations of the vertex), but it is not necessary, as it depends on the grain generator (it is like scanning a sky filled with pulsating stars). Also note that density does not depend only on spatial distribution of the grains on the map, but mostly on graph structures.

Also the note approach can be simulated in GeoGraphy, i.e. GeoGraphy allows composers to work with vertices that represents events of greater complexity than grains. Consider the graph in Figure 7. In the example, vertices 1-9 represent a subgraph in which all the possible paths starting from 1-2-3 have a duration in range between 143 and 221 milliseconds, which can be considered as the duration of the whole subgraph as a higher-level note event. The whole graph can be considered as a two note cell, the first resulting from subgraph 1-9 and the second one from subgraph

10-17. The vertices 1-9 are grouped so that four sets of increasing edge and vertex labels emerges (a-d): with opportune tuning, this kind of relation can be used to model a typical percussive noisy attack with subsequent periodicity and decay.

The advantages of the GeoGraphy model for the composer rely mostly upon the symbolic approach, in that each graph structure represents a set of relations between vertices, which can be thought as *objets sonores* ([24]). Sound objects as defined by Schaeffer are symbolic objects encoding sonic properties apt to be used in compositional practice. In this sense, they are considered meaningful-to-the-ear objects, while physical dimensions need a continuous monitoring to be musically relevant. Sound objects have been placed at the analogous of the phonological level in language, as sets of relevant sonic features (*acoulogie* in [24], sonology in Laske's term, see [25] for a general discussion). GeoGraphy is intended to be a model to structure sequences of well-fit sound events (*objets convenables*), of which notes in traditional composition, but also microsounds, are special cases.

#### 4. CONCLUSIONS

This paper has presented a formal system for the algorithmic control of composition based on granular synthesis. The system features a grain level, that organizes grains into a graph structure, and a spatial level, that distributes the graphs in specific locations of a space. The composition process is the design of a trajectory in the space, appropriately interpreted to control a number of parameters of physical and musical relevance. We have also seen how the system can be viewed as a generalization of the current approaches to composition with granular synthesis.

#### 3. REFERENCES

- [1] De Poli, G., "Granular Representations of Musical Signals: Overview", in De Poli, G. - Piccialli, A. - Roads, C. (eds.), Representations of Musical Signals, The MIT Press, Cambridge (Mass.)-London, 1991.
- [2] Roads, C., Microsound, The MIT Press, Cambridge (Mass.)-London, 2001.
- [3] Roads, C., "Introduction to Granular Synthesis", Computer Music Journal, vol. 12, n. 2, 1988.
- [4] Roads, C., The computer music tutorial, The MIT Press, Cambridge (Mass.)-London, 1996.
- [5] Xenakis, I., Formalized Music. Thought and Mathematics in Composition, Bloomington, Indiana, 1971 (rev. ed. Pendragon Press, Stuyvesant (NY), 1991).
- [6] Wishart, T., Audible Design. A Plain and Easy Introduction to Practical Sound Composition, Orpheus the Pantomime, 1994.

- [7] Truax, B., "Composing with Real-Time Granular Synthesis", *Perspectives of New Music*, vol. 28, n. 2, 1990.
- [8] Jones, D. L. – Parks, T. W., "Generation and Combination of Grains for Music Synthesis", *Computer Music Journal*, vol. 12, n. 2, 1988.
- [9] Lee, A. S.C., "Granular Synthesis in Csound", in Boulanger, R., *The Csound Book. Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming*, The MIT Press, Cambridge (Mass.)-London, 2000.
- [10] Truax, B., "Real-Time Granular Synthesis with a Digital Signal Processor", *Computer Music Journal*, vol. 12, n. 2, 1988.
- [11] Roads, C., "Asynchronous Granular Synthesis", in De Poli, G. - Piccialli, A. - Roads, C. (ed.), *Representations of Musical Signals*, The MIT Press, Cambridge (Mass.)-London, 1991.
- [12] Roads, C., "Granular Synthesis of Sound", in Roads, C.-Strawn, J., *Foundations of Computer Music*, The MIT Press, Cambridge (Mass.)-London, 1985.
- [13] Giordani, E., "GSC4 – Sintesi granulare per Csound", in Bianchini, R. – Cipriani, A., *Il Suono Virtuale. Csound per PC e Mac. Teoria e Pratica*, ConTempo, Roma, 1998.
- [14] Bartetzki, A., "Csound Score Generation and Granular Synthesis with Cmask" (on-line paper: <http://www.kgw.tu-berlin.de/~abart/CMaskPaper/cmask-article.html>, visited on Nov 27th, 2002).
- [15] Diestel, R., *Graph Theory*, Springer (electronic ed.), New York, 2000.
- [16] Valle, A., "Del *Gran Paese*, or: Some brief remarks on space", unpublished paper, 2002.
- [17] Plomp, R., *Aspects of Tone Sensation. A Psychophysical Study*, Academic Press, London, 1976.
- [18] Wessel, D., "Timbre Space as a Musical Control Structure", *Computer Music Journal*, vol. 3, n. 2, 1979.
- [19] Rasch, R.A. - Plomp, R., "The Perception of Musical Tones", in Deutsch, D., *The Psychology of Music*, Academic Press, Orlando, 1982.
- [20] Slawson, W., *Sound Color*, University of California Press, Berkeley-Los Angeles-London, 1985.
- [21] Risset, J. C. – Wessel, D., "Exploration of Timbre by Analysis and Synthesis", in Deutsch, D., *The Psychology of Music*, Academic Press, Orlando, 1982.
- [22] Lerdahl, F., "Les hiérarchies de timbres", in Barrière, J.B. (ed.), *Le timbre. Métaphore pour la composition*, Christian Bourgois-IRCAM, Paris, 1991 (French transl. of "Timbral Hierarchies", *Contemporary Music Review*, vol. 2, n. 1, 1987).
- [23] Shepard, R. N., "Structural Representations of Musical Pitch", in Deutsch, D., *The Psychology of Music*, Academic Press, Orlando, 1982.
- [24] Schaeffer, P., *Traité des objets musicaux*, Seuil, Paris, 1966.
- [25] Bel, B., "Symbolic and Sonic Representations of Sound-Objet Structure", in Balaban, M., Ebcioglu, K., Laske, O. (eds.), *Understanding Music with AI: Perspectives on Music Cognition*, The AAI Press/The MIT Press, Cambridge (Mass.)-Menlo Park-London, 1992.